Estimates for representation numbers of binary quadratic forms and Apollonian circle packings

Radu Toma

Abstract

Fix a primitive, positive definite binary quadratic form g with integer coefficients. We prove asymptotic formulas for sums of the form $\sum r_g(n)^\beta$ and $\sum r_g^*(n)^\beta$, where $\beta \geq 0$ and $r_g(n)$, resp. $r_g^*(n)$, denote the number of inequivalent representations, resp. proper inequivalent representations, of n by g. These estimates generalize a previous result by Blomer and Granville (2006) by allowing for non-fundamental discriminants and also clarify some details in the proof of the positive density conjecture for integral Apollonian circle packings by Bourgain and Fuchs (2011).

Keywords: Quadratic forms, Non-fundamental discriminants, Proper representations, Apollonian circle packings

1. Introduction

If $g(x,y) = ax^2+bxy+cy^2$ is a positive definite binary quadratic form with integer coefficients, how many numbers smaller than some positive X does g represent and with what multiplicity? For forms[1] with fundamental discriminants, Blomer and Granville [1] answered this question by giving estimates uniform in the discriminant. One remarkable application of their results is an important step in the proof of the positive density conjecture for integral Apollonian circle packings by Bourgain and Fuchs [3].

To recall the estimates in [1], let the form g have discriminant $-D < 0$ and let $N_g(X)$ be the number of distinct integers smaller than X that are represented by g. It turns out that the function $N_g(X)$ changes behaviour depending essentially on the relation between X and the discriminant. Blomer and Granville identify three ranges and give estimates for $N_g(X)$ in each of them. This article focuses on the first range, when $(\log X)^{2+\varepsilon} \leq D = o(X)$, which is covered by Theorem 2 in [1]. Given these conditions, the theorem implies that

$$N_g(X) \sim C_g \frac{X}{\sqrt{D}}, \tag{1}$$

with an explicit constant C_g depending on the form. The bounds on the error term are also explicit in their dependence on the discriminant, which makes this result particularly useful when working with families of quadratic forms with varying discriminant. This idea was used in [3] to count curvatures in Apollonian circle packings, as explained in the next paragraphs. This will serve as additional motivation for taking a closer look at the details of Blomer and Granville's theory.

[1] Throughout this article, the words *form* and *quadratic form* should be taken to mean *positive definite primitive quadratic form in two variables with integer coefficients.*

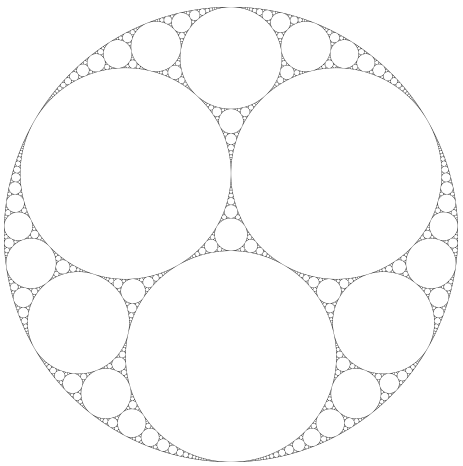

Figure 1: An Apollonian circle packing[2]

An Apollonian circle packing (ACP) is constructed by starting with three mutually tangent circles. A theorem of the ancient Greek geometer Apollonius of Perga states that there are precisely two other circles tangent to all three given ones. Intuitively, these are the big circle enclosing them all and the smaller circle inscribed in the triangular interstice formed between the three. This new configuration of five circles gives rise to new interstices, in which we can inscribe unique circles by the same argument. Filling all gaps with circles like this inductively, we obtain an Apollonian circle packing as the limit of this process (see Figure 1).

By definition, all circles in an integral ACP have integer curvatures (inverses of the radii). There is an abundance of integral circle packings and there has been a recent increase in interest towards the arithmetic properties of such packings (see [5] or [6] for an overview). One natural question that arose is whether the set of curvatures in an arbitrary (bounded) integral ACP has positive density inside the natural numbers. Bourgain and Fuchs [3] showed that the answer is affirmative. One of the main ingredients of their proof is an ingenious application of [1, Theorem 2], though, strictly speaking, in a generalized version.

More precisely, for an integral Apollonian circle packing P, let $\kappa(P, X)$ be the number of distinct integers up to X occurring as curvatures of circles in P. Theorem 1.2 of [3] states that there exists a constant c depending on the packing P, such that

$$\kappa(P, X) \gg_c X,$$

for X big enough. The key observation in the proof is that the set of integers appearing as curvatures in the packing contains the integers properly represented by a family of shifted binary quadratic forms

$$f_a(x, y) - a,$$

indexed by certain integers a. This family needs to be chosen carefully in order to apply the asymptotic in (1) and obtain lower bounds on the cardinality of this union of represented integers.

[2]This image is reproduced from `https://commons.wikimedia.org/wiki/File:Apollonian_gasket.svg` under CC BY-SA 4.0 licence.

Indeed, Bourgain and Fuchs apply the first step of the inclusion-exclusion principle to obtain

$$\begin{aligned}\kappa(P,X) &\geq \#\bigcup_a \{n \in \mathbb{N} : n \leq X, n \text{ represented by } f_a - a\} \\ &\geq \sum_a \#\{n \leq X, n \text{ represented by } f_a - a\} \\ &\quad - \sum_{a \neq a'} \#\{n \leq X, n \text{ represented by both } f_a - a \text{ and } f_{a'} - a'\}.\end{aligned}$$

The problem of bounding the first sum from below almost readily lends itself to Blomer and Granville's results, whilst an upper bound for the second sum requires different methods. These tasks become rather delicate, as one needs to achieve the right balance between the two bounds. In particular, it is essential to find an appropriate set of indices a over which to sum:[3] a bigger index set gives a better lower bound in the first problem, but weakens the upper bound for the intersections. Sharpening the methods in the second problem to control this trade-off is a large part of the labour in [3].

The first sum is bounded from below in [3] (specifically in the proof of Lemma 3.1) using the estimates provided in [1, Theorem 2]. However, in their article, Blomer and Granville count representations for forms with fundamental discriminants. They mention that, with some extra work, the results generalize to non-fundamental discriminants, but the details do no seem to be in print. On the other hand, Bourgain and Fuchs count *proper* representations for forms f_a having discriminant $-4a^2$, which is fundamental if and only if the integer a has absolute value 1. Now, since we are interested in a lower bound, requiring representations to be proper is a non-trivial condition. Moreover, given a positive X, the index a ranges between $(\log X)^2$ and $(\log X)^3$, implying that the form f_a has a non-fundamental discriminant virtually always. Thus, Bourgain and Fuchs implicitly use a generalized version of [1, Theorem 2], although they give an alternative proof for the special case they need in the appendix of [3].

This article intends to make the aforementioned generalizations of [1, Theorem 2] explicit. This contributes in itself to the theory of binary quadratic forms and is of independent interest, striving towards a completion of Blomer and Granville's useful theory. It conveniently comes with the extra benefit of clarifying some details of the proof in [3]. The main theorem of this article is a version of [1, Theorem 2] in which non-fundamental discriminants are allowed.

We denote by $\mathfrak{C}(D)$ the class group of positive definite forms with discriminant $-D$ and let w_D be the number of automorphisms of such forms. Inside $\mathfrak{C}(D)$ we find the subgroup $\mathfrak{G}(D)$ of ambiguous classes, i.e. classes with order at most 2. We may decompose the discriminant as $-D = D_0 f^2$, where D_0 is a fundamental discriminant and f is the conductor of D. For each divisor d of f there is a homomorphism $\theta_d : \mathfrak{C}(D) \longrightarrow \mathfrak{C}(D/d^2)$, which will be made precise in Section 3. The results of Sun and Williams [7] make this homomorphism an important tool for us. Using this notation and the convention that $r_g(n)^0 = 0$ if $r_g(n) = 0$ and $r_g(n)^0 = 1$ otherwise, the main theorem is stated below. The main term looks relatively complicated, but as we shall discuss in Section 5.1 this is inevitable, at least in general. The error term is smaller than the main term as soon as $(\log X)^{2+\varepsilon} \leq |D| = o(X)$.

Theorem 1. *Let g be a binary quadratic form of discriminant $-D = D_0 f^2 < 0$ with conductor f. For each divisor d of f, let $a_{\theta_d(g)}$ be the smallest positive integer represented by $\theta_d(g)$ and let*

[3] N.B. These integers a come with their own restrictions (they are a subset of the curvatures), and one must first prove that this set is large enough. This already requires a good bound going towards the positive density conjecture.

$u_{\theta_d(g)}$ be the smallest positive integer coprime to f/d that can be represented by some form in the coset $\theta_d(g)\mathfrak{G}(D/d^2)$. For any $\beta \geq 0$ we have:

$$\sum_{n \leq x} r_g(n)^\beta = \frac{\pi x}{\sqrt{D}} \cdot \frac{2}{w_D^\beta} \sum_{d|f} \left[w_{D/d^2}^{\beta-1} \cdot \frac{\varphi(f/d)}{f} \left(1 + \frac{2^{\beta-1}-1}{u_{\theta_d(g)}} \right) \right] + E_\beta(x, D),$$

where

$$E_\beta(x, D) \ll \begin{cases} \displaystyle\sum_{d|f} \frac{2^{\omega(f/d)}}{d} \sqrt{\frac{x}{a_{\theta_d(g)}}} + \tau(f^2) + \tau(D) \left(\frac{x \log x}{D} + \frac{x}{D^{3/4}} \right), & 0 \leq \beta \leq 2, \\ \displaystyle\sum_{d|f} \frac{2^{\omega(f/d)}}{d} \sqrt{\frac{x}{a_{\theta_d(g)}}} + \tau(f^2) + \tau(D) \frac{x (\log x)^{(2/q)(2^{(\beta-2)q+1}-1)+1}}{D^{(3/4)(1-1/q)}}, & \beta > 2, \end{cases}$$

for any real $q > 1$, where $\tau(D)$ denotes the number of divisors of D. The implied constants depend at most on β and q.

Similar results are given for counting *proper* representations in Section 4, though not as general. Nevertheless, these results suffice for the application to the proof of the positive density conjecture for integral Apollonian circle packings, as shown at the end of Section 4.

2. Representing integers coprime to the conductor

At the most basic level, the main tool used by Blomer and Granville to prove [1, Theorem 2] is the well-known correspondence between classes of quadratic forms and ideal classes. Forms with fundamental discriminants correspond to ideals in the ring of integers of quadratic number fields. For non-fundamental discriminants, we must regard more general, i.e. non-maximal, quadratic orders and restrict the notion of ideals to proper ideals. In both cases, this correspondence induces a bijection between representations of integers and ideals with certain norms, as in the next lemma (see [2, Chap. 2, Sect. 7, Theorem 5]).

Lemma 2. *Let g be a form with discriminant $-D < 0$ and let $\mathcal{O}_D$ be the quadratic order of discriminant $-D$. There is a bijection between inequivalent solutions to $g(x, y) = n > 0$ and proper $\mathcal{O}_D$-ideals of norm n in a class C_g.*

In essence, Blomer and Granville decompose the ideals corresponding to representations into two factors and estimate the possibilities for each factor. Finding this decomposition relies heavily on prime ideal factorization, this being the main difficulty in the generalization to non-fundamental discriminants. Indeed, non-maximal quadratic orders do not have unique prime factorization of ideals and the best way to recover this property is to restrict to the ideals coprime to the conductor (see [4, Sect. 7.C]). Therefore, we first prove a weaker version of the main theorem following the proof of Theorem 2 in [1], restricting to integers coprime to the conductor.

Theorem 3. *For a binary quadratic form g having discriminant $-D = D_0 f_D^2 < 0$ with conductor f_D, let a_g be the smallest positive integer that is represented by g, and let u_g be the smallest positive integer coprime to f_D that can be represented by some form in the coset $g\mathfrak{G}(D)$. For any $\beta \geq 0$ we have:*

$$\sum_{\substack{n \leq x \\ (n, f_D)=1}} r_g(n)^\beta = \frac{\varphi(f_D)}{f_D} \cdot \frac{2}{w_D} \left(1 + \frac{2^{\beta-1}-1}{u_g} \right) \frac{\pi x}{\sqrt{D}} + E_\beta(x, D),$$

where

$$E_\beta(x,D) \ll \begin{cases} 2^{\omega(f)}\left(1+\sqrt{\frac{x}{a_g}}\right)+2^{\omega(D)}\left(\frac{x\log x}{D}+\frac{x}{D^{3/4}}\right), & 0\le\beta\le 2,\\ 2^{\omega(f)}\left(1+\sqrt{\frac{x}{a_g}}\right)+2^{\omega(D)}\frac{x(\log x)^{(2/q)(2^{(\beta-2)q+1}-1)+1}}{D^{(3/4)(1-1/q)}}, & \beta>2, \end{cases}$$

for any real $q>1$, where $\omega(D)$ denotes the number of prime divisors of D. The implied constants depend at most on β and q.

Proof. We may take the proof in [1, pp. 279–282] and simply require all integers, as well as ideals and indices of sums, to be coprime to the conductor. The correspondence in Lemma 2 is preserved under this restriction, since an ideal is coprime to the conductor if and only if its norm is so (see [4, Lemma 7.18, Chap. 2]). Moreover, prime ideal factorization is recovered, so that all decompositions and facts derived from this property are valid in this case as well.

To give a sketch of the proof, recall first that a primitive ideal is an ideal that is not divisible by any rational integer other than 1. Now, Blomer and Granville observe that a pair of different ideals of norm n in the class C_g (corresponding to inequivalent representations as in Lemma 2) can be decomposed as $\mathfrak{b}\mathfrak{c}$ and $\mathfrak{b}\bar{\mathfrak{c}}$, where $\mathfrak{c}$ is an ideal in some ambiguous class G, such that $\mathfrak{c}\neq\bar{\mathfrak{c}}$, and $\mathfrak{b}$ is a primitive ideal coprime to D in the class C_gG. This decomposition is done algorithmically and uses prime ideal factorization and ramification theory[4], which is available if n is coprime to f.

Since the decomposition is not explicitly done in [1], we provide the details here. Note that two ideals $\mathfrak{a}_1,\mathfrak{a}_2\in C_g$ with norm n coprime to the conductor have the form $\mathfrak{a}_1=\prod\mathfrak{p}_i\prod\mathfrak{q}_j\prod(r_k)$ and $\mathfrak{a}_2=\prod\mathfrak{p}_i\prod\mathfrak{s}_j\prod(r_k)$, where $\{\mathfrak{p}_i\}_i$ are prime ideals over the ramified primes, $\{\mathfrak{q}_j,\mathfrak{s}_j\}_j$ correspond to the split primes, and $\{r_k\}_k$ are the inert primes. Thus, the two outer products $\prod\mathfrak{p}_i$ and $\prod(r_k)$ are identical in both factorizations, respectively, since $N(\mathfrak{a}_1)=N(\mathfrak{a}_2)$. We want $\mathfrak{b}$ to be primitive and coprime to D, so that the product $\prod\mathfrak{p}_i\prod(r_k)$ needs to divide $\mathfrak{c}$.

Now let $Q:=\{q \text{ prime} : q\mid n,(\frac{D_0}{q})=1\}$ and for each $q\in Q$ choose a prime ideal denoted by $\mathfrak{q}$ that contains q. Then the middle products are of the form

$$\prod\mathfrak{q}_j=\prod_{q\in Q}\mathfrak{q}^{i_q}\bar{\mathfrak{q}}^{j_q},\quad \prod\mathfrak{s}_j=\prod_{q\in Q}\mathfrak{q}^{k_q}\bar{\mathfrak{q}}^{l_q}, \tag{2}$$

where $i_q+j_q=k_q+l_q$ for all $q\in Q$. To see how to construct $\mathfrak{b}$ and $\mathfrak{c}$, let us assume without loss of generality that $i_q=\min(i_q,j_q,k_q,l_q)$ for a prime $q\in Q$. Then the factor in (2) corresponding to q is of the form

$$\mathfrak{q}^{i_q}\bar{\mathfrak{q}}^{j_q}=(q^{i_q})\bar{\mathfrak{q}}^{j_q-i_q},\quad \mathfrak{q}^{k_q}\bar{\mathfrak{q}}^{l_q}=(q^{i_q})\mathfrak{q}^{k_q-i_q}\bar{\mathfrak{q}}^{l_q-i_q},$$

using that $\mathfrak{q}\bar{\mathfrak{q}}=q$. Since we want $\mathfrak{b}\in\mathfrak{A}$, the principal ideal (q^{i_q}) should divide $\mathfrak{c}$. Denoting $\mathfrak{b}_q=\bar{\mathfrak{q}}^{l_q-i_q}\subset\mathcal{O}_D$ and $\mathfrak{c}_q=(q^{i_q})\bar{\mathfrak{q}}^{m}\subset\mathcal{O}_D$ where $m=j_q-i_q-(l_q-i_q)=k_q-i_q\in\mathbb{N}_0$, we have

$$\mathfrak{q}^{i_q}\bar{\mathfrak{q}}^{j_q}=\mathfrak{b}_q\mathfrak{c}_q,\quad \mathfrak{q}^{k_q}\bar{\mathfrak{q}}^{l_q}=\mathfrak{b}_q\overline{\mathfrak{c}_q}.$$

Defining $\mathfrak{b}_q$ and $\mathfrak{c}_q$ analogously for all $q\in Q$ and denoting

$$\mathfrak{b}=\prod_{q\in Q}\mathfrak{b}_q,\quad \mathfrak{c}=\prod_i\mathfrak{p}_i\prod_{q\in Q}\mathfrak{c}_q\prod_k(r_k),$$

[4]Since the isomorphism in [4, Prop. 7.20] preserves norms and commutes with complex conjugation, we may use the same properties of ramified, inert and split primes as in the case of the maximal order.

we find that $\mathfrak{a}_1 = \mathfrak{b}\mathfrak{c}$ and $\mathfrak{a}_2 = \mathfrak{b}\bar{\mathfrak{c}}$. By construction, we see that $\mathfrak{b}$ is primitive and coprime to D. Moreover, $\mathfrak{a}_1 \neq \mathfrak{a}_2$ implies that $\mathfrak{a}_1\mathfrak{b}^{-1} = \mathfrak{c} \neq \bar{\mathfrak{c}} = \mathfrak{a}_2\mathfrak{b}^{-1}$. Since $\mathfrak{a}_1$ and $\mathfrak{a}_2$ are in the same class, there exists $\xi \in \mathcal{O}_D$ such that $\xi\mathcal{O}_D = \mathfrak{a}_1\overline{\mathfrak{a}_2} = N(\mathfrak{b})\mathfrak{c}^2$, and it follows that $\mathfrak{c}$ is in an ambiguous class.

In view of this decomposition, let us denote by $\mathfrak{G}(D)$ the set of ambiguous ideal classes with discriminant $-D$ and for each class $G \in \mathfrak{G}(D)$ write $\mathfrak{X}_G := \{\mathfrak{c} \in G \mid \mathfrak{c} \neq \bar{\mathfrak{c}}\}$. We denote the set of primitive ideals coprime to D by $\mathfrak{A}$. Out of all the possible factors $\mathfrak{b}$ we can distinguish one particular ideal $\mathfrak{u}$, which is the ideal in some class $C_G G_0$ of the coset $C_g\mathfrak{G}(D)$ having the smallest norm $N\mathfrak{u} = u_g$ coprime to the conductor. This ideal $\mathfrak{u}$ is in fact the only possibility for the factor $\mathfrak{b}$ if $n \leq (D/4)^{1/4}$. This and more is the content of Lemma 5.1 in [1], which gives estimates for the cardinalities

$$\#\{\mathfrak{c} \in \mathfrak{X}_G \mid N\mathfrak{c} \leq X\} \quad \text{and} \quad \#\{\mathfrak{b} \in C_g G \cap \mathfrak{A} \mid \mathfrak{b} \neq \mathfrak{u}, N\mathfrak{b} \leq X\}, \tag{3}$$

for each $G \in \mathfrak{G}(D)$.

Our proof now needs estimates for the sets in (3), where we additionally require that the norms of the ideals be coprime to the conductor. The same bounds as in [1, Lemma 5.1] suffice in our case as well and may be proven in the same way. For the first set recall that each ambiguous class contains a reduced form with a simple shape (see [4, Lemma 3.10]). This allows us to crudely count representations and, most importantly, exclude enough of them which do not correspond to ideals not equal to their conjugates and get a good bound. We exclude the same ideals as Blomer and Granville, since these are either not coprime to the conductor or are equal to their conjugates as in the original proof. The estimates for primitive ideals are done the same way as in [1]. The proofs again implicitly use prime ideal factorization and ramification theory.

Next, we can approximate the sum we are interested in by one for which we can find a good asymptotic. For this we define

$$A_1(n) := \#\{\mathfrak{a} \in C_g \mid N\mathfrak{a} = n, a \notin u\mathfrak{X}_{G_0}\} \quad \text{and} \quad A_2(n) := \#\{\mathfrak{a} \in C_g \mid N\mathfrak{a} = n, a \in u\mathfrak{X}_{G_0}\}.$$

Further we define

$$B := \{\mathfrak{c} \in G_0 \mid \mathfrak{c} = \bar{\mathfrak{c}}, N(\mathfrak{c}) \leq x/u_g, (N(\mathfrak{c}), f) = 1\} \quad \text{and} \quad r_g^*(n,\beta) := A_1(n) + 2^{\beta-1}A_2(n).$$

A generalization of Lemma 3.1 in [1], given below as Lemma 4, provides the asymptotic

$$|\{\mathfrak{a} \in C_g \mid N\mathfrak{a} \leq x, (N\mathfrak{a}, f) = 1\}| = \frac{2}{w_D} \cdot \frac{\varphi(f)}{f} \cdot \frac{2\pi x}{\sqrt{D}} + O\left(2^{\omega(f)}\left(1 + \sqrt{\frac{x}{a}}\right)\right),$$

using the correspondence between (inequivalent) representations and ideals. The same computations made by Blomer and Granville in [1, Eq. (5.1)] show that

$$\begin{aligned}
\sum_{\substack{n \leq x \\ (n,f)=1}} r_g^*(n,\beta) &= \sum_{\substack{n \leq x \\ (n,f)=1}} (A_1(n) + A_2(n)) + (2^{\beta-1} - 1)\left(|B| + \sum_{\substack{n \leq x \\ (n,f)=1}} A_2(n)\right) + O_\beta(|B|) \\
&= |\{\mathfrak{a} \in C_g \mid N\mathfrak{a} \leq x, (N\mathfrak{a}, f) = 1\}| + \\
&\quad + (2^{\beta-1} - 1) \cdot |\{\mathfrak{c} \in G_0 \mid N\mathfrak{c} \leq x/u_g, (N\mathfrak{c}, f) = 1\}| + O\left(\sqrt{\frac{x}{a_g}}\right) \\
&= \frac{\varphi(f)}{f} \cdot \frac{2}{w_D}\left(1 + \frac{2^{\beta-1} - 1}{u_g}\right)\frac{\pi x}{\sqrt{D}} + O\left(2^{\omega(f)}\left(1 + \sqrt{\frac{x}{a_g}}\right)\right).
\end{aligned}$$

Approximating the sum over $r_g(n)^\beta$ by the sum over $r_g^*(n,\beta)$ gives an error for which the bounds differ depending on whether $0 \le \beta \le 2$ or $\beta > 2$. In both cases, we can bound the sum

$$\sum_{\substack{1\le n\le x\\(n,f)=1}} |r_g(n)^\beta - r_g^*(n,\beta)|$$

by a sum over the positive integers coprime to f up to X containing terms essentially of the form $r_g(n)^2 - r_g^*(n,2)$. This difference is equal to the number of pairs of different ideals in C_g with norm n that, decomposed as $(\mathfrak{bc}, \mathfrak{b}\bar{\mathfrak{c}})$, have $\mathfrak{b} \neq \mathfrak{u}$. To get an upper bound, we multiply the two estimates from [1, Lemma 5.1] together, and our proof here continues almost identically to the one given by Blomer and Granville. We merely need to apply the bound $|\mathfrak{G}| \ll 2^{\omega(D)}$ instead of the one used in [1], that is $|\mathfrak{G}| \ll \tau(D)$, which is not optimal any more for non-fundamental discriminants.

Broadly speaking, the only ingredient in the proof of Blomer and Granville which does not generalize immediately is the asymptotic in [1, Lemma 3.1], where a more precise approach is needed (see Lemma 4). □

Lemma 4. *If g is a positive definite primitive quadratic form of discriminant $-D < 0$ with conductor f and a is the smallest integer represented by g, then*

$$\#\{(m,n) \in \mathbb{Z}^2 \mid g(m,n) \le x, (g(m,n), f) = 1\} = \frac{\varphi(f)}{f} \cdot \frac{2\pi x}{\sqrt{D}} + O\left(2^{\omega(f)}\left(1 + \sqrt{\frac{x}{a}}\right)\right).$$

The implied constant is absolute.

Proof. We may first assume that $g(x_1,x_2) = ax_1^2 + b'x_1x_2 + c'x_2$ is reduced. Let $f = f_a\tilde{f}$ be a decomposition of the conductor such that $(a,\tilde{f}) = 1$ and all primes dividing f_a divide a as well. The second part of Lemma 2.1 in [7] shows[5] that we can assume g to have the form

$$g(x_1,x_2) = ax_1^2 + b\tilde{f}x_1x_2 + c\tilde{f}^2x_2^2.$$

Since the coefficient a is represented by g, Theorem 3.2 in [7] (given in (4) below) implies that $(a,f^2) = (a,f_a^2) = k^2$ for some $k \in \mathbb{Z}$. By our construction of f_a we find that all primes p dividing f_a must satisfy $p^2 \mid a$ and, since $-D = b^2 - 4ac$ and g is primitive, p also divides b and does not divide c. Now let $g(x_1,x_2) = n$ for some $n \in \mathbb{N}$ and suppose p is a prime such that $p \mid (n,f)$. If $p \mid f_a$, then $(c\tilde{f}^2, p) = 1$, $p \mid b$ and $p^2 \mid a$, so that from Lemma 2.2 of [7] it follows that $p \mid x_2$. If $p \mid \tilde{f}$, then $p \mid x_1$, again by Lemma 2.2 of [7]. Therefore, $(g(x_1,x_2),f) = 1$ if and only if $(x_1,\tilde{f}) = 1$ and $(x_2,f_a) = 1$.

Equipped with this criterion for coprimality, we may now follow the rest of the proof from [1]. The two coprimality conditions lead to the slightly weaker error term and to the factor

$$\frac{\varphi(f)}{f} = \frac{\varphi(f_a)}{f_a} \cdot \frac{\varphi(\tilde{f})}{\tilde{f}}$$

in front of the main term. □

Remark 5. The factor $2/w_D$ in the main term of Theorem 3 arises by switching from counting all representations in Lemma 4 to only inequivalent representations in the theorem. If $-D < -4$, then this factor can be neglected, since $w_D = 2$.

[5] Note that in the proof of [7, Lemma 2.1, (ii)] we can start with any first coefficient a as long as $(a,m) = 1$, in the notation of [7]. In our case, m is equal to $\tilde{f}$.

3. Proof of the main theorem

To prove Theorem 1 we reduce the general case to the coprime case, where we can apply Theorem 3. For this we need the following facts, proved by Sun and Williams in [7].

Let the quadratic form g have discriminant $-D$ with conductor f and let C be the equivalence class of g. Lemma 2.1 of [7] states that for any integer m dividing f we can find a form in C that has coefficients (a, mb, m^2c), with a, b and c integers. The map θ_m sends the class C to the class of forms equivalent to (a, b, c).[6] Note that the form $ax^2 + bxy + cy^2$ now has discriminant $-D/m^2$ and conductor f/m. By abuse of notation, we will also use the form g itself as the argument of θ_m.

For a positive integer n, let $R(g, n)$ be the number of integer solutions to the equation $n = g(x, y)$. Theorem 3.2 of [7] asserts that

$$R(g,n) = \begin{cases} 0 & \text{if } (n, f^2) \text{ is not a square,} \\ R(\theta_m(g), n/m^2) & \text{if } (n, f^2) = m^2 \text{ for } m \in \mathbb{N}. \end{cases} \tag{4}$$

Note that, in the second case, n/m^2 is coprime to the conductor of $\theta_m(g)$, which is f/m.

We may now rearrange the sum we are interested in by divisors of the conductor. We have

$$\begin{aligned} \sum_{n\le x} r_g(n)^\beta &= \sum_{n\le x} \frac{1}{w_D^\beta} R(g,n)^\beta = \sum_{d\mid f} \sum_{\substack{n\le x \\ (n,f^2)=d^2}} \frac{1}{w_D^\beta} R(g,n)^\beta \\ &= \sum_{d\mid f} \left(\frac{w_{D/d^2}}{w_D}\right)^\beta \sum_{\substack{k\le x/d^2 \\ (k,(f/d)^2)=1}} r_{\theta_d(g)}(k)^\beta. \end{aligned}$$

Applying Theorem 3, we first compute the main term as

$$\begin{aligned} &\sum_{d\mid f} \left(\frac{w_{D/d^2}}{w_D}\right)^\beta \cdot \frac{\varphi(f_{D/d^2})}{f_{D/d^2}} \cdot \frac{2}{w_{D/d^2}} \left(1 + \frac{2^{\beta-1}-1}{u_{\theta_d(g)}}\right) \frac{\pi x/d^2}{\sqrt{D/d^2}} \\ &= \frac{\pi x}{\sqrt{D}} \cdot \frac{2}{w_D^\beta} \sum_{d\mid f} \left[w_{D/d^2}^{\beta-1} \cdot \frac{\varphi(f/d)}{f} \left(1 + \frac{2^{\beta-1}-1}{u_{\theta_d(g)}}\right)\right]. \end{aligned}$$

Next, for the error term we merely need the convolution identity $\sum_{d\mid f} 2^{\omega(f/d)} = \tau(f^2)$. This directly shows that

$$\sum_{d\mid f} 2^{\omega(f/d)} \left(1 + \sqrt{\frac{x}{d^2 a_{\theta_d(g)}}}\right) = \tau(f^2) + \sum_{d\mid f} \frac{2^{\omega(f/d)}}{d} \sqrt{\frac{x}{a_{\theta_d(g)}}}.$$

Another straightforward application of the convolution identity and recalling the decomposition $-D = D_0 f^2$, where D_0 is a fundamental discriminant, quickly proves that

$$\sum_{d\mid f} 2^{\omega(D/d^2)} \le \tau(D). \tag{5}$$

[6] This map is well defined and is denoted by ϕ_m in Definition 2.1 of [7].

We finally bound $\log(x/d^2)$ by $\log x$ for each divisor d of the conductor[7]. Thus we have

$$\sum_{d|f} 2^{\omega(D/d^2)} \left(\frac{(x/d^2)\log(x/d^2)}{D/d^2} + \frac{x/d^2}{(D/d^2)^{3/4}} \right) \ll \tau(D) \left(\frac{x \log x}{D} + \frac{x}{D^{3/4}} \right),$$

which gives the error term in the case $0 \leq \beta \leq 2$. The estimate for $\beta > 2$ is analogous.

Remark 6. Note that the w_D and w_{D/d^2} factors can be often ignored as in Remark 5, but there is indeed a contribution if the fundamental discriminant D_0 is -3 or -4.

4. Proper representations and the ACP estimate

A quadratic form g properly represents an integer n if there are integers x and y such that $g(x, y) = n$ and $(x, y) = 1$. We denote the number of proper representations by $R^*(g, n)$ and define $r_g^*(n) = R^*(g, n)/w_D$, where D is the discriminant of g.

It is easy to see that

$$R(g, n) = \sum_{d^2|n} R^*(g, n/d^2)$$

and applying Möbius inversion yields that

$$R^*(g, n) = \sum_{d^2|n} \mu(d) R(g, n/d^2).$$

This identity combined with Lemma 3.1 in [1] and Lemma 4 swiftly produces analogues of these lemmata.

Lemma 7. *For a binary quadratic form g with discriminant $-D < 0$ and conductor f, let a be the smallest integer represented by g. Then we have the asymptotics*

$$\sum_{n \leq x} R^*(g, n) = \frac{1}{\zeta(2)} \cdot \frac{2\pi x}{\sqrt{D}} + O\left(\sqrt{\frac{x}{a}} \log x \right)$$

and

$$\sum_{\substack{1 \leq n \leq x \\ (n,f)=1}} R^*(g, n) = \frac{1}{\zeta(2)} \cdot \frac{\phi(f)}{f} \cdot \frac{2\pi x}{\sqrt{D}} + O\left(2^{\omega(f)} \left(\sqrt{x} + \sqrt{\frac{x}{a}} \log x \right) \right).$$

Now notice that, in Lemma 2, proper representations correspond to primitive ideals. Indeed, recall that the isomorphism between representations of n and proper ideals of the order $\mathcal{O}_D$ with discriminant $-D < 0$ having norm n is given by

$$(x, y) \mapsto (xa + ya\tau)\mathfrak{a}^{-1},$$

where $\tau = (-b + \sqrt{-D})/2a$ and $\mathfrak{a}$ is the ideal generated by a and $a\tau$ as a $\mathbb{Z}$-module. If $\gcd(x, y) > 1$, then $(xa + ya\tau)\mathfrak{a}^{-1} = \gcd(x, y)(\frac{x}{\gcd(x,y)}a + \frac{y}{\gcd(x,y)}a\tau)\mathfrak{a}^{-1}$ is not primitive. Conversely, if $kI = (xa + ya\tau)\mathfrak{a}^{-1}$ for some $\mathcal{O}_D$-ideal I and $k \in \mathbb{Z}$, then $kI\mathfrak{a} = (xa + ya\tau)\mathcal{O}_D$. Thus $(xa + ya\tau) = k(x'a + y'a\tau)$, which is easily seen to imply that $k \mid x$ and $k \mid y$.

We may restrict all ideals to primitive ideals in the proof of Theorem 3 and obtain its analogue using Lemma 7.

[7] In Theorem 3 the $\log x$ terms in the error do not appear if $x < 1$, so that we correctly bound $\log(x/d^2)$ by $\log x$ when $x/d^2 \geq 1$ and simply bound 0 instead of $\log(x/d^2)$ by $\log x$ otherwise.

Corollary 8. *For a given binary quadratic form g with discriminant $-D = D_0 f_D^2 < 0$ and conductor f_D, let a_g be the smallest positive integer that is represented by g, and let u_g be the smallest positive integer coprime to f_D that can be represented by some form in the coset $g\mathfrak{G}(D)$. For any $\beta \geq 0$ we have:*

$$\sum_{\substack{n \leq x \\ (n, f_D)=1}} r_g^*(n)^\beta = \frac{1}{\zeta(2)} \cdot \frac{\varphi(f_D)}{f_D} \cdot \frac{2}{w_D} \left(1 + \frac{2^{\beta-1}-1}{u_g}\right) \frac{\pi x}{\sqrt{D}} + E_\beta(x, D),$$

where

$$E_\beta(x, D) \ll \begin{cases} 2^{\omega(f)} \left(\sqrt{x} + \sqrt{\frac{x}{a_g}} \log x\right) + 2^{\omega(D)} \left(\frac{x \log x}{D} + \frac{x}{D^{3/4}}\right), & 0 \leq \beta \leq 2, \\ 2^{\omega(f)} \left(\sqrt{x} + \sqrt{\frac{x}{a_g}} \log x\right) + 2^{\omega(D)} \frac{x (\log x)^{(2/q)(2^{(\beta-2)q+1}-1)+1}}{D^{(3/4)(1-1/q)}}, & \beta > 2, \end{cases}$$

for any real $q > 1$. The implied constants depend at most on β and q.

Remark 9. The proof simply uses the same bounds as in the proof of Theorem 3, since $R^*(g, n) \leq R(g, n)$. Note here that a_g and u_g are properly represented in virtue of their definition.

This result cannot be generalized to an analogue of the main theorem as in Section 3, since the reduction identity (4) fails when restricting to proper representations. See Section 5.2 for a counter example.

Turning to Apollonian circle packings, we now use Theorem 1 and Lemma 7 to complete the proof of Lemma 3.1 in [3]. The objects we are concerned with here are quadratic forms f_a of discriminant $-4a^2$ and the integers they properly represent. More precisely, we need the estimate

$$\sum_{n \leq X} R^*(f_a, n)^0 \gg \frac{X}{a}, \tag{6}$$

where a is an integer between $(\log X)^2$ and $(\log X)^3$. This is the only step in the proof that needs further clarification.

To obtain (6) we use the Cauchy-Schwarz inequality to reduce the problem to known cases. We have

$$\sum_{n \leq X} R^*(f_a, n)^0 \geq \frac{\left(\sum_{n \leq X} R^*(f_a, n)\right)^2}{\sum_{n \leq X} R^*(f_a, n)^2} \geq \frac{\left(\sum_{n \leq X} R^*(f_a, n)\right)^2}{\sum_{n \leq X} R(f_a, n)^2}.$$

This is similar to the application of Cauchy-Schwarz in the appendix of [3].

To bound the denominator, note first that the discriminant $-4a^2$ lies in absolute value between $(\log X)^4$ and $(\log X)^6$, which implies that the error term $E_2(X, -4a^2)$ in Theorem 1 is smaller than the main term. Indeed, the first summand in the error is

$$\sum_{d|a} \frac{2^{\omega(a/d)}}{d} \sqrt{\frac{X}{a_{\theta_d(g)}}} \leq \sqrt{X} \sum_{d|a} 2^{\omega(a/d)} = \tau(D^2)\sqrt{X} \ll X^{1/2+\varepsilon}.$$

Next, the last summand can be bounded by

$$D^\varepsilon \left(\frac{X \log X}{D} + \frac{X}{D^{3/4}}\right) \ll \frac{X}{D^{3/4-\varepsilon}},$$

where we used that $\log X \ll D^{1/4}$. Finally, the main term is bounded from above by a constant times X/a, since we may bound the w_{D/d^2} factors trivially and

$$\sum_{d|f} \frac{\varphi(f/d)}{f} \left(1 + \frac{2^{2-1}-1}{u_{\theta_d(g)}}\right) \leq 2 \sum_{d|f} \frac{\varphi(f/d)}{f} = 2,$$

where we used the crude bound $u_{\theta_d(f_a)} \geq 1$.

The lower bound for the numerator is provided by Lemma 7, since the error term is smaller than the main term, by an argument analogous to the above. The desired estimate (6) now follows and the implied constant does not depend on a.

5. Remarks

5.1. The constant in the main term of the main theorem

In comparison to the original theorem of Blomer and Granville, the generalized Theorem 1 has a more complicated main term. One would like to simplify the sum over the divisors by finding some relations between the numbers $u_{\theta_d(g)}$, since computing these is far from trivial and can be quite expensive. Thankfully, if g is in an ambiguous class, then $u_g = 1$, since the coset $g\mathfrak{G}(D)$ includes the class of the principal form, which obviously represents 1. Analogously, for all divisors d of the conductor we have $u_{\theta_d(g)} = 1$, since the images of g under the maps θ_d are ambiguous classes as well. Indeed, Theorem 2.1 in [7] states that θ_d is in fact a surjective homomorphism, preserving the property of having order 1 or 2. Therefore, assuming $D_0 < -4$ so that all w_{D/d^2} are 2 for convenience, the factor after $\pi X/\sqrt{D}$ in the main term simplifies to

$$(1 + 2^{\beta-1}) \sum_{d|f} \frac{\varphi(f/d)}{f} = (1 + 2^{\beta-1}).$$

Since all $u_{\theta_d(g)}$ are at least 1, this is the maximal value of this factor.

In contrast, for non-ambiguous classes of forms we cannot expect this kind of cancelling. Under the assumption above, simplifying the factor comes down to understanding the sum

$$\sum_{d|f} \frac{\varphi(f/d)}{f u_{\theta_d(g)}}.$$

The values of $u_{\theta_d(g)}$ usually vary with d and the sum does not necessarily equal a fraction of the form $1/u$, as one would naively try to generalize Theorem 2 of [1]. Table 1, computed using the computer algebra system *Magma*, shows a few examples of non-ambiguous forms of discriminant $-8575 = -7 \cdot (5 \cdot 7)^2$. Here, the number $\tilde{u}_g$ denotes the smallest integer represented by the coset $g\mathfrak{G}(D)$ without the condition of coprimality to the conductor.

g	$\tilde{u}_g$	u_g	$u_{\theta_5(g)}$	$u_{\theta_7(g)}$	$u_{\theta_{35}(g)}$	$\sum \frac{\varphi(f/d)}{f u_{\theta_d(g)}}$
$[2, 1, 1072]$	2	2	2	2	1	$\frac{59}{70}$
$[25, 5, 86]$	25	86	1	2	1	$\frac{381}{430}$
$[49, 35, 50]$	49	53	2	1	1	$\frac{1697}{1855}$

Table 1: Computed examples for forms of discriminant -8575

One possible idea for understanding this behaviour is to use the reduction identity (4) and try to relate the numbers u_g and $u_{\theta_d(g)}$. This strategy would at least require the homomorphisms θ_d to map ambiguous classes to ambiguous classes surjectively, so that the cosets $g\mathfrak{G}(D)$ and $\theta_d(g)\mathfrak{G}(D/d^2)$ are correlated. Unfortunately, this is not necessarily true for even discriminants. For instance, there are exactly two ambiguous classes of discriminant $-256 = -4 \cdot 8^2$, namely the classes of the forms with coefficients $(17, -4, 4)$ and $(1, 0, 64)$.[8] Now notice that, if $[a, b, c]$ denotes the class of the form $ax^2 + bxy + cy^2$, then

$$\theta_2([1, 0, 64]) = [1, 0, 16] \quad \text{and} \quad \theta_2([17, -4, 4]) = [17, -2, 1] = [1, 0, 16].$$

On the other hand, there are two ambiguous classes of discriminant -64, so that the map θ_2 restricted to ambiguous classes is not surjective in this case.

5.2. Considerations for a main theorem for proper representations

Applying the same strategy as in Section 3 to prove an analogue of the main theorem for proper representations is not directly possible. Indeed, Theorem 3.2 of [7], i.e. equation (4), does not hold when restricting to proper representations. For a counter-example consider the form $g(x, y) = x^2 + 36y^2$ with discriminant $-4 \cdot (2 \cdot 3)^2 = -144$. Its image under θ_3 is the form $\tilde{g}(x, y) = x^2 + 4y^2$ with discriminant $-4 \cdot 2^2$. The number $n = 37$ has, up to equivalence, a single representation, namely $\tilde{g}(1, 3) = \tilde{g}(-1, -3) = 37$, which is proper. On the other hand, the number $m = n \cdot 3^2 = 333$ has only one equivalence class of representations, namely $g(3, 3) = g(-3, -3) = 333$, which is not proper.

It seems thus necessary to have a better understanding of primitive ideals and their factorizations in non-maximal quadratic orders to achieve the desired result in a fashion similar to the proof in [1].

Acknowledgements

I am most grateful to my supervisor, Prof. Valentin Blomer, for introducing me to these topics and for his kind and valuable guidance. I also thank the reviewer for their constructive and important input, as well as my friends Christian Bernert, Vlad Crișan, Jan Gundelach, and Sabyasachi Mukherjee for their feedback and suggestions.

References

[1] Blomer, V., Granville, A., 2006. Estimates for representation numbers of quadratic forms. Duke Mathematical Journal 135, 261–302.

[2] Borevich, Z.I., Shafarevich, I.R., 1966. Number theory. Academic Press.

[3] Bourgain, J., Fuchs, E., 2011. A proof of the positive density conjecture for integer apollonian circle packings. J. Amer. Math. Soc. 24, 945–967.

[4] Cox, D.A., 2013. Primes of the form $x^2 + ny^2$. 2 ed., Wiley.

[5] Fuchs, E., 2013. Counting problems in apollonian packings. Bulletin of the American Mathematical Society 50, 229–266.

[8] To compute the number of ambiguous forms we use [4, Prop. 3.11] and to find explicit forms we use [4, Lemma 3.10].

[6] Graham, R.L., Lagarias, J.C., Mallows, C.L., Wilks, A.R., Yan, C.H., 2003. Apollonian circle packings: Number theory. J. Number Theory 100, 1–45.

[7] Sun, Z.H., Williams, K.S., 2006. On the number of representations of n by $ax^2 + bxy + cy^2$. Acta Arithmetica 122, 101–171.